REMARKS

By the present amendment, the abstract have been amended in the manner suggested by the Examiner. Claim 6 has been amended to clarify features thereof, as discussed below. The specification does not require any correction since the proper spacing is provided between the words, as illustrated in US Publication 2008/0231272.

With respect to rejection of claims 1-8 under 35 USC 112, first paragraph, as failing to comply with the enablement requirement because the term "pseudo inverse matrix" is not described or defined in the specification, applicants submit that the definition of the term "pseudo inverse matrix" is well known, as disclosed, e.g., in Wikipedia (copy attached). That is to say, since the definition of this term is general and well known, there is no need to recite the definition of this term in the specification. Thus, claims 1-8 are considered to be in compliance with 35 USC 112, first paragraph.

By the above amendment, claim 6 has been amended to clarify the features of the present invention, in particular to provide the structural relationship between the sliding window recited and rest of the elements or steps recited in claim 6. Claim 6 has been amended to recite (in relevant part):"...the control unit... (3) employs a sliding window to obtain the plurality of images, and (4) performs the scanning of the k-space at intervals of n echoes and suppresses artifacts by a low-pass filter after image reconstruction". The limitations (3) and (4) are described by way of example on page 15, lines 9-16. For example, in the present invention, in order to suppress artifacts occurring on an edge of an image due to under-sampling, data to which the sliding window has not been applied and which is acquired by displacing scanning lines for fear scanning lines for adjoining frames may be superimposed on each

other is employed. The temporal filter (low-pass filter) is applied to the time-sequential data, whereby artifacts derived from under-sampling can be suppressed. Applicants submit that such features as now recited in claim 6 are in compliance with 35 USC 112, second paragraph, and patentably distinguish over the cited art as will become clear from the following discussion.

With respect to the rejection of claims 6-8 under 35 USC 112, second paragraph, as noted above, independent claim 6 has been amended to be in compliance with 35 USC 112, second paragraph. Accordingly, dependent claims 7 and 8 are considered in compliance with 35 USC 112, second paragraph.

As to the rejection of claim 1-5 under 35 USC 102(b) as being anticipated by Tasaka et al. P.N. 6,515,477; the rejection of claims 6-8 under 35 USC 102(b) as being anticipated by Meyer et al. P.N. 5,485, 086, such rejections are traversed insofar as they are applicable to the present claims, and reconsideration and withdrawal of the rejections are respectfully requested.

Independent claim 1 (taken as an example) recites (in relevant part):"...the processing unit (1) divides the image echoes and reference echoes into a plurality of groups, (2) uses the reference echo and image echoes preceding and succeeding the reference echo to calculate an estimation coefficient, and (3) uses the estimation coefficient to estimate unmeasured echoes lying among the image echoes in the k-space" (emphasis added). These features are described, by way of example only, in lines 1-25, on page 11, in lines 1-2, on page 12, in lines 8-13, on page 13, in lines 5-11, on page 14 of the specification and illustrated in FIGS. 4-7.

The present invention provides a magnetic resonance imaging technology for efficiently suppressing artifacts in radial scanning. In the present invention, the processing unit 104 illustrated in FIG. 3 measures seventy-two echoes. The

positions of echoes in the k-spaces are indicated with solid lines as illustrated in FIG. 4. In FIG. 4, among echoes indicated with solid lines, an echo 213 indicated with a boldface line is a reference echo. The echo is used to calculate an estimation coefficient, whereby unmeasured echoes 212 indicated with dot lines are estimated. The sum total of echoes including estimated ones comes to one hundred and twenty-eight echoes. The procedure will be described in conjunction with the flowchart of FIG. 6. To begin with, one hundred and twenty-eight echoes including the unmeasured echoes 212 are divided into eight groups one of which includes seventeen echoes (step 401). As shown in FIG. 5, marginal echoes of each group are identical to marginal echoes of two adjoining groups. In FIG. 5, an echo 214 is included as a marginal echo in each of the first group 301 and second group 302. The reference echo 213 is measured to be included in the middle of each group. Thereafter, an estimation coefficient A = [a1, a2] to be used to estimate the unmeasured echoes 212 included in each group is calculated according to the expression (1) below (step 402).

$$A=RS - \qquad (1)$$

where S equals [S1, S2] T (a T denotes a transposed matrix), S-denotes a pseudo inverse matrix, and S1 and S2 (column vectors) are echoes adjacent to the reference echo R (row vector). Thereafter, the estimation coefficient A is used to estimate an unmeasured echo according to the expression (2) below (step 403).

$$Su=AS' \qquad (2)$$

where Su denotes an unmeasured echo, and S' equals [S'1, S'2] T where S'1 and S'2 denote echoes adjacent to Su. By performing the above processing, each echo is equally divided into Np parts as indicated with dashed lines 303 in FIG. 5. FIG. 5 is concerned with a case where Np equals 3. When an echo is divided, a more highly

precise estimate is obtained than the one obtained when one echo is used as it is. Finally, measured echoes and unmeasured echoes that are estimated as mentioned above are combined and gridded. Thereafter, inverse Fourier transform is performed in order to reconstruct an image (step 404).

As mentioned above, according to the present invention, part of unmeasured echoes is measured as a reference. An estimation coefficient is calculated using echoes adjoining the reference, and the unmeasured echoes are estimated using the estimation coefficient. Consequently, since an echo measured as the reference is one of the unmeasured echoes, an imaging time hardly increases. Moreover, since the reference is used to calculate the estimation coefficient, unmeasured echoes can be highly and precisely estimated compared to when they are estimated by simple interpolation without using the reference. This results in a magnetic resonance imager capable of suppressing artifacts with little extension of the imaging time.

On the other hand, Tasaka et al. discloses in FIGS. 9-17 a radial trajectory formation in sequence. Tasaka discloses in col. 6, lines 60-68:

> first, data collection is performed on a first trajectory 1 along the kx axis, as shown in FIG. 9. Next, data is collected on a second trajectory 2 along the ky axis, as shown in FIG. 10. The trajectories 1 and 2 are orthogonal to each other, and have an angular difference of π/2. Thus, the k-space is divided into four quadrants. Then, data are collected respectively on trajectories 3 and 4 each of which has a direction bisecting the angular difference π/2 between the trajectories 1 and 2, as shown in FIG. 11. Thus, the angular difference between the trajectories becomes π/4.

Furthermore, Takasa discloses that when the data collection is performed, every trajectory has a different direction than the immediately preceding one by an angle of or near π /2. Thus, time phases of motion of the imaged object 300 are approximately randomly distributed over the k-space. Therefore, regularity among the view data associated with the motion directivity is diluted. Thus, Tasaka

discloses a signal collection method by collecting magnetic resonance signals along a plurality of linear radial trajectories passing through center of k-space. However, Tasaka fails to disclose a processing unit that divides the image echoes and reference echoes into a plurality of groups. Tosaka also fails to disclose a processing unit that calculates an estimation coefficient by using the reference echo and image echo preceding and succeeding the reference echo. Tosaka also fails to disclose a processing unit that estimates unmeasured echoes lying among the image echoes in the k-space.

Furthermore, Tasaka discloses making Cartesian (grid) data from radial data by interpolation method called gridding. The interpolation by gridding sums radial data with the weight response to distance from the interpolation (grid) point. Thus, the weight is predetermined and is not dependent on the image. In the present invention disclosed in claim 1, the estimation coefficient is calculated by using image echoes. Therefore, as it was stated above, the interpolation suitable for each image can be precisely performed as compared with interpolation method disclosed in Tasaka. Thus, Tasaka fails to disclose or teach:"...the processing unit (1) divides the image echoes and reference echoes into a plurality of groups, (2) uses the reference echo and image echoes preceding and succeeding the reference echo to calculate an estimation coefficient, and (3) uses the estimation coefficient to estimate unmeasured echoes lying among the image echoes in the k-space" (emphasis added), as disclosed in independent claim 1 of the application. Thus, claims 1-5 are not anticipated by Tasaka, and should be considered allowable thereover.

With respect to the rejection of claims 6-8, independent claim 6 (taken as an example) has been amended to recite (in relevant part):"...the control unit (1) detects the nuclear magnetic resonance signal by radially scanning a k-space, (2) produces

a plurality of images, (3) employs a sliding window to obtain the plurality of images, and (4) performs the scanning of the k-space at intervals of n echoes and suppresses artifacts by a low-pass filter after image reconstruction" (emphasis added). These features are described, by way of example only, in lines 5-25, on page 16, in lines 1-15, on page 17 of the specification and illustrated in FIGS. 8-11. FIG. 10 shows the results of imaging performed according to the present invention. For example, FIG. 10(C) shows an image produced by scanning the k-space at intervals of three echoes, applying the temporal filter (low-pass filter), and suppressing artifacts derived from under-sampling. Not only streak artifacts attributable to a gap between adjoining data items but also artifacts appearing on an edge of an image are seen being suppressed. One example of the present invention has been described on the assumption that the k-space is scanned at intervals of three echoes. At interval of how many echoes the k-space should be scanned is determined using the expression (3) below.

$$(N+1)=Ne/2 \qquad (3)$$

where N denotes the number of echoes N at intervals of which the k-space is scanned, and Ne denotes a magnification by which a frame rate is increased by sharing echoes. The expression (3) is drawn out as described below. Assuming that a frame rate is multiplied by Ne by sharing echoes, a gap between adjoining data items varies in cycles of Ne/F (where F denotes a frame rate). Accordingly, a streak artifact having a frequency F/Ne occurs. When the k-space is scanned at intervals of N echoes, the number of echoes to be scanned while a scanning line makes one turn is decreased to 1/(N+1). Consequently, the frequency of a change in the gap between adjoining data items is multiplied by (N+1). Accordingly, the frequency of a streak artifact comes to F/Ne x (N+1). The expression (3) is drawn out under the

condition of (F/Ne x (N+1) = F/2) that the frequency of a streak artifact should be high (Nyquist rate F/2), where Ne is magnification ratio of frame rate by sliding window.

Meyer et al. discloses in FIGS. 6A and 6B the temporal filtering of the K-space signals for fluoroscopic imaging. In FIGS. 6A and 6B the moving window 40 moves with time to combine the latest three detected signals for the fluoroscopic image. Thus, in Meyer the contrast of the image can be improved by making the weight of the latest signal large. In contrast, in the invention recited in claim 6, a low path filter is used in the time direction of each frame image after image reconstruction.

In the constitution which measures at intervals of N-echoes, the claimed invention, as it was stated above, can eliminate artifacts with the low-pass filter by shifting the frequency of the artifacts to high frequency region under the expression (N+1)=Ne/2. Thus, Meyer fails to disclose or teach: "...the control unit (1) detects the nuclear magnetic resonance signal by radially scanning a k-space, (2) produces a plurality of images, (3) employs a sliding window to obtain the plurality of images, and (4) performs the scanning of the k-space at intervals of n echoes and suppresses artifacts by a low-pass filter after image reconstruction" (emphasis added), as recited in independent claim 6 of the application. Thus, claims 6-8 are not anticipated by Meyer, and should be considered allowable thereover.

In view of above amendments and remarks, applicants submit that claims present in this application should now be in condition for allowance and issuance of an action favorable nature is courteously solicited.

If the Examiner believes that there are any other points which may be clarified or otherwise disposed of either by telephone discussion or by personal interview, the

Examiner is invited to contact Applicants' undersigned attorney at the number indicated below.

To the extent necessary, Applicants petition for an extension of time under 37 CFR 1.136. Please charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, to the Antonelli, Terry, Stout & Kraus, LLP Deposit Account No. 01-2135 (Docket No. 520.46403X00), and please credit any excess fees to such deposit account.

Respectfully submitted,
ANTONELLI, TERRY, STOUT & KRAUS, LLP

/Alan E Schiavelli/
Alan E Schiavelli
Registration No. 32,087

IS/AES
1300 North Seventeenth Street, Suite 1800
Arlington, Virginia 22209
Telephone: (703) 312-6600
Facsimile: (703) 312-6666

Generalized inverse

From Wikipedia, the free encyclopedia

(Redirected from Pseudoinverse)
"Pseudoinverse" redirects here. For the Moore-Penrose pseudoinverse, sometimes referred to as "the pseudoinverse", see Moore-Penrose pseudoinverse.

In mathematics, a **generalized inverse** or **pseudoinverse** of a matrix A is a matrix that has some properties of the inverse matrix of A but not necessarily all of them. The term "the pseudoinverse" commonly means the Moore-Penrose pseudoinverse.

The purpose of constructing a generalized inverse is to obtain a matrix that can serve as the inverse in some sense for a wider class of matrices than invertible ones. Typically, the generalized inverse exists for an arbitrary matrix, and when a matrix has an inverse, then its inverse and the generalized inverse are the same. Some generalized inverses can be defined in any mathematical structure that involves associative multiplication, that is, in a semigroup.

Contents

Types of generalized inverses

The various kinds of generalized inverses include

- one-sided inverse, that is left inverse and right inverse
- Drazin inverse
- Group inverse
- Bott–Duffin inverse
- Moore-Penrose pseudoinverse

See also

- Inverse element

References

- Bing Zheng and R. B. Bapat, *Generalized inverse A(2)T,S and a rank equation*, Applied Mathematics and Computation 155 (2004) 407-415 DOI 10.1016/S0096-3003(03)00786-0 (http://dx.doi.org/10.1016/S0096-3003(03)00786-0)
- S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Dover 1991 ISBN 978-0486666938
- Adi Ben-Israel and Thomas N.E. Greville, *Generalized inverses. Theory and applications*. 2nd ed. New York, NY: Springer, 2003. ISBN 0-387-00293-6 Zbl 1026.15004 (http://www.zentralblatt-math.org/zmath/en/search/?q=an:1026.15004&format=complete)
- C. Radhakrishna Rao and Sujit Kumar Mitra, *Generalized Inverse of Matrices and its Applications*, John Wiley & Sons New York, 1971, 240 p., ISBN 0-471-70821-6

External links

- 15A09 (http://www.ams.org/msc/15-xx.html) Matrix inversion, generalized inverses in Mathematics Subject Classification, MathSciNet search (http://www.ams.org/mathscinet/search/publications.html?pg4=AUCN&s4=&co4=AND&pg5=TI&s5=&co5=AND&pg6=PC&s6=15A09&co6=AND&pg7=ALLF&s7=&co7=AND&Submit=Search&dr=all&yrop=eq&arg3=&y
- Pseudo-Inverse (Not Moore-Penrose) (http://mjollnir.com/matrix/demo.html)
- googlevideo - lecture at MIT dealing with Pseudomatrices (http://video.google.com/videoplay?docid=-8273560482088448841)

Retrieved from "http://en.wikipedia.org/wiki/Generalized_inverse"
Categories: Matrices | Mathematical disambiguation

Moore-Penrose pseudoinverse

From Wikipedia, the free encyclopedia

In mathematics, and in particular linear algebra, the **pseudoinverse** A^+ of an $m \times n$ matrix A is a generalization of the inverse matrix.[1] More precisely, this article talks about the **Moore-Penrose pseudoinverse**, which was independently described by E. H. Moore[2] in 1920 and Roger Penrose[3] in 1955. Earlier, Fredholm had introduced the concept of a pseudoinverse of integral operators in 1903. The term generalized inverse is sometimes used as a synonym for pseudoinverse.

A common use of the pseudoinverse is to compute a 'best fit' (least squares) solution to a system of linear equations that lacks a unique solution (see below under Applications). The pseudoinverse is defined and unique for all matrices whose entries are real or complex numbers. It can be computed using the singular value decomposition.

Contents

Definition

The pseudoinverse A^+ of an m-by-n matrix A (whose entries can be real or complex numbers) is defined as the unique n-by-m matrix satisfying all of the following four criteria:[3][4]

1. $AA^+A = A$ (AA^+ need not be the general identity matrix, but it maps all column vectors of A to themselves);

2. $A^+AA^+ = A^+$ (A^+ is a weak inverse for the multiplicative semigroup);
3. $(AA^+)^* = AA^+$ (AA^+ is Hermitian); and
4. $(A^+A)^* = A^+A$ (A^+A is also Hermitian).

Here M^* is the hermitian transpose (also called conjugate transpose) of a matrix M. For matrices whose elements are real numbers instead of complex numbers, $M^* = M^T$.

Properties

Proofs for some of these relations may be found on the proofs page.

- The pseudoinverse exists and is unique: for any matrix A, there is precisely one matrix A^+ that satisfies the four properties of the definition.[4] If A has real entries, then so does A^+; if A has rational entries, then so does A^+.
- If the matrix A is invertible, the pseudoinverse and the inverse coincide: $A^+ = A^{-1}$.[5]:243
- The pseudoinverse of a zero matrix is its transpose.
- The pseudoinverse of the pseudoinverse is the original matrix: $(A^+)^+ = A$.[5]:245
- Pseudoinversion commutes with transposition, conjugation, and taking the conjugate transpose:[5]:245

$$\begin{aligned}(A^T)^+ &= (A^+)^T,\\ \overline{A}^+ &= \overline{A^+},\\ (A^*)^+ &= (A^+)^*.\end{aligned}$$

- The pseudoinverse of a scalar multiple of A is the reciprocal multiple of A^+:

  $(\alpha A)^+ = \alpha^{-1}A^+$ for $\alpha \neq 0$.

- If A and B are such that the product AB is defined and
  - A has orthonormal columns ($A^*A = I$, unitarity is a special case), or
  - B has orthonormal rows ($BB^* = I$), or
  - A is of full column rank, and B is of full row rank,

  then

  $(AB)^+ = B^+A^+$.

  The third case does not cover the first two; an orthonormal (including non-square) matrix must be of full rank, but otherwise there is no assumption made on the matrix it multiplies.

- AA^+ is the orthogonal projector onto the range of A (the space spanned by the column vectors of A); A^+A is the orthogonal projector onto the range of A^*, which agrees with the orthogonal complement of the kernel (null space) of A; $(1 - A^+A)$ is the orthogonal projector onto the kernel (null space) of A.[4]
- The kernel of A^+ is the orthogonal complement of the range of A; the image of A^+ is the orthogonal complement of the kernel of A.
- If the pseudoinverse of A^*A is already known, it may be used to compute A^+:

$$A^+ = (A^* A)^+ A^* .$$

- Likewise, if $(AA^*)^+$ is already known:

$$A^+ = A^* (AA^*)^+ .$$

- The pseudoinverse can be computed via a limiting process:

$$A^+ = \lim_{\delta \to 0} (A^* A + \delta I)^{-1} A^* = \lim_{\delta \to 0} A^* (AA^* + \delta I)^{-1}$$

(see Tikhonov regularization). These limits exist even if $(AA^*)^{-1}$ and $(A^* A)^{-1}$ do not exist.[4]:263

- In contrast to ordinary matrix inversion, the process of taking pseudoinverses is not continuous: if the sequence (A_n) converges to the matrix A (in the maximum norm or Frobenius norm, say), then $(A_n)^+$ need not converge to A^+.

Identities

The following identities can be used to cancel certain subexpressions or expand expressions involving pseudoinverses. Proofs for these properties can be found in the proofs subpage.

$$\begin{aligned}
A^+ &= A^+ A^{+*} A^* \\
A^+ &= A^* A^{+*} A^+ \\
A &= A^{+*} A^* A \\
A &= A A^* A^{+*} \\
A^* &= A^* A A^+ \\
A^* &= A^+ A A^*
\end{aligned}$$

Special cases

Orthonormal columns or rows

If A has orthonormal columns ($A^* A = I$) or orthonormal rows ($AA^* = I$), then $A^+ = A^*$.

Full rank

If the **columns** of A are linearly independent, then $A^* A$ is invertible. In this case, an explicit formula is:[1]

$$A^+ = (A^* A)^{-1} A^* .$$

It follows that A^+ is a left inverse of A: $A^+ A = I$.

If the **rows** of A are linearly independent, then AA^* is invertible. In this case, an explicit formula is:

$$A^+ = A^* (AA^*)^{-1}.$$

It follows that A^+ is a right inverse of A: $AA^+ = I$.

If both **columns and rows** are linearly independent (that is, for square regular matrices), the pseudoinverse is just the inverse:

$$A^+ = A^{-1}.$$

Scalars and vectors

It is also possible to define a pseudoinverse for scalars and vectors. This amounts to treating these as matrices. The pseudoinverse of a scalar x is zero if x is zero and the reciprocal of x otherwise:

$$x^+ = \begin{cases} 0, & \text{if } x = 0; \\ x^{-1}, & \text{otherwise.} \end{cases}$$

The pseudoinverse of the null vector is the transposed null vector. The pseudoinverse of a non-null vector is the conjugate transposed vector divided by its squared magnitude:

$$x^+ = \begin{cases} 0^T, & \text{if } x = 0; \\ \frac{x^*}{x^* x}, & \text{otherwise.} \end{cases}$$

For a proof, simply check that these definitions meet the defining criteria for the pseudoinverse.

Circulant matrices

For a Circulant matrix C the singular value decomposition is given by the Fourier transform, that is the singular values are the Fourier coefficients. Let $\mathcal{F}$ be the Fourier matrix, then

$$C = \mathcal{F} \cdot \Sigma \cdot \mathcal{F}^*$$
$$C^+ = \mathcal{F} \cdot \Sigma^+ \cdot \mathcal{F}^*$$ [6]

Block matrices

Optimized approaches exist for calculating the pseudoinverse of block structured matrices.

Finding the pseudoinverse of a matrix

Using regular inverses

Let k be the rank of a $m \times n$ matrix A. Then A can be decomposed as $A = BC$, where B is a $m \times k$ matrix and C is a $k \times n$ matrix. Then

$$A^+ = C^* (CC^*)^{-1} (B^* B)^{-1} B^*.$$

If A has full row rank, so that $k = m$, then B can be chosen to be the identity matrix and the formula reduces to $A^+ = A^* (AA^*)^{-1}$. Similarly, if A has full column rank (that is, $k = n$), we have $A^+ = (A^* A)^{-1} A^*$.

The QR method

However, computing the product AA^* or $A^* A$ or its inverse explicitly is unnecessary and only causes additional rounding errors and computational cost. Consider first the case when A is full column rank. Then

all that is needed is the Cholesky decomposition

$$A^* A = R^* R,$$

where R is an upper triangular matrix. Multiplication by the inverse is then done easily by solving a system with multiple right-hand-side,

$$M = (A^*A)^{-1}N \quad \Leftrightarrow \quad (A^*A)M = N \quad \Leftrightarrow \quad R^*RM = N$$

by back substitution. To compute the Cholesky decomposition without forming $A^* A$ explicitly, use the QR decomposition of A, $A = QR$ where Q is a unitary matrix, $Q^* Q = QQ^* = I$, and R is upper triangular. Then

$$A^*A = (QR)^*(QR) = R^*Q^*QR = R^*IR = R^*R,$$

so R is the Cholesky factor of $A^* A$. The case of full row rank is obtained by swapping A and A^*.

The general case and the SVD method

A computationally simpler and more accurate way to get the pseudoinverse is by using the singular value decomposition.[1][4][7] If $A = U\Sigma V^*$ is the singular value decomposition of A, then $A^+ = V\Sigma^+ U^*$. For a diagonal matrix such as Σ, we get the pseudoinverse by taking the reciprocal of each non-zero element on the diagonal, and leaving the zeros in place. In numerical computation, only elements larger than some small tolerance are taken to be nonzero, and the others are replaced by zeros. For example, in the Matlab function `pinv`, the tolerance is taken to be $t = \varepsilon \cdot \max(m,n) \cdot \max(\Sigma)$, where ε is the machine epsilon.

The cost of this method is dominated by the cost of computing the SVD, which is several times higher than matrix-matrix multiplication, if a state-of-the art implementation (such as that of LAPACK) is used.

The above procedure shows why taking the pseudoinverse is not a continuous operation: if the original matrix A has a singular value 0 (a diagonal entry of the matrix Σ above), then modifying A slightly may turn this zero into a tiny positive number, thereby affecting the pseudoinverse dramatically as we now have to take the reciprocal of a tiny number.

Updating the pseudoinverse

For the cases where A has full row or column rank, and the inverse of the correlation matrix (AA^* for A with full row rank or $A^* A$ for full column rank) is already known, the pseudoinverse for matrices related to A can be computed by applying the Sherman–Morrison–Woodbury formula to update the inverse of the correlation matrix, which may need less work. In particular, if the related matrix differs from the original one by only a changed, added or deleted row or column, incremental algorithms[8] exist that exploit the relationship.

Similarly, it is possible to update the Cholesky factor when a row or column is added, without creating the inverse of the correlation matrix explicitly. However, updating the pseudoinverse in the general rank-deficient case is much more complicated.[9][10]

Software libraries

High quality implementations of SVD, QR, and back substitution are available in standard libraries, such as LAPACK. Writing one's own implementation of SVD is a major programming project that requires a significant numerical expertise. In special circumstances, such as parallel computing or embedded computing, however, alternative implementations by QR or even the use of an explicit inverse might be preferable, and custom implementations may be unavoidable.

Applications

See also: Hat matrix

The pseudoinverse provides a least squares solution to a system of linear equations.[11] Given a system of linear equations

$$Ax = b,$$

in general we cannot always expect to find a vector x which will solve the system; even if there exists such a solution vector, then it may not be unique. We can however always ask for a vector x that brings Ax "as close as possible" to b, i.e. a vector that minimizes the Euclidean norm

$$\|Ax - b\|^2.$$

If there are several such vectors x, we could ask for the one among them with the smallest Euclidean norm. Thus formulated, the problem has a unique solution, given by the pseudoinverse:

$$x = A^+ b.$$

This description suggests the following geometric construction of the pseudoinverse of an $m \times n$ matrix A. To find $A^+ b$ for given b in $\mathbf{R}^m$, first project b orthogonally onto the range of A, finding a point $p(b)$ in the range. Then form $A^{-1}(\{p(b)\})$, i.e. find those vectors in $\mathbf{R}^n$ that A sends to $p(b)$. This will be an affine subspace of $\mathbf{R}^n$ parallel to the kernel of A. The element of this subspace that has the smallest length (i.e. is closest to the origin) is the answer $A^+ b$ we are looking for. It can be found by taking an arbitrary member of $A^{-1}(\{p(b)\})$ and projecting it orthogonally onto the orthogonal complement of the kernel of A.

Using the pseudoinverse and a matrix norm, one can define a condition number for any matrix:

$$\operatorname{cond}(A) = \|A\| \|A^+\|$$

A large condition number implies that the problem of finding least-squares solutions to the corresponding system of linear equations is ill-conditioned in the sense that small errors in the entries of A can lead to huge errors in the entries of the solution.[12]

Generalization

In order to solve more general least-squares problems, one could try to define Moore-Penrose pseudoinverses for all continuous linear operator $A : H_1 \to H_2$ between two Hilbert spaces H_1 and H_2, using the same four conditions as in our definition above. It turns out that not every continuous linear operator has a continuous linear pseudo-inverse in this sense.[12] Those that do are precisely the ones whose range is closed in H_2.

See also

- Hat matrix

References

1. ^ [a] [b] [c] Ben-Israel, Adi; Thomas N.E. Greville (2003). *Generalized Inverses*. Springer-Verlag. ISBN 0-387-00293-6.
2. ^ Moore, E. H. (1920). "On the reciprocal of the general algebraic matrix". *Bulletin of the American Mathematical Society* **26**: 394–395.
3. ^ [a] [b] Penrose, Roger (1955). "A generalized inverse for matrices". *Proceedings of the Cambridge Philosophical Society* **51**: 406–413.
4. ^ [a] [b] [c] [d] [e] Golub, Gene H.; Charles F. Van Loan (1996). *Matrix computations* (3rd edition ed.). Baltimore: Johns Hopkins. pp. 257–258. ISBN 0-8018-5414-8.
5. ^ [a] [b] [c] Stoer, Josef; Bulirsch, Roland (2002), *Introduction to Numerical Analysis* (3rd ed.), Berlin, New York: Springer-Verlag, ISBN 978-0-387-95452-3.
6. ^ Stallings, W. T. (1972). "The Pseudoinverse of an r-Circulant Matrix". *Proceedings of the American Mathematical Society* **34**: 385–388. doi:10.2307/2038377 (http://dx.doi.org/10.2307%2F2038377) .
7. ^ Linear Systems & Pseudo-Inverse (http://www.uwlax.edu/faculty/will/svd/systems/index.html)
8. ^ Tino Gramß (1992) (printed dissertation). *Worterkennung mit einem künstlichen neuronalen Netzwerk*. Georg-August-Universität zu Göttingen.
9. ^ Meyer, Carl D., Jr. Generalized inverses and ranks of block matrices. SIAM J. Appl. Math. 25 (1973), 597—602
10. ^ Meyer, Carl D., Jr. Generalized inversion of modified matrices. SIAM J. Appl. Math. 24 (1973), 315—323
11. ^ Penrose, Roger (1956). "On best approximate solution of linear matrix equations". *Proceedings of the Cambridge Philosophical Society* **52**: 17–19.
12. ^ [a] [b] Roland Hagen, Steffen Roch, Bernd Silbermann. *C*-algebras and Numerical Analysis*, CRC Press, 2001. Section 2.1.2.

External links

- Pseudoinverse on PlanetMath (http://planetmath.org/encyclopedia/Pseudoinverse.html)
- Moore-Penrose inverse (http://planetmath.org/?op=getobj&from=objects&id=6067) on PlanetMath
- *Eric W. Weisstein, Pseudoinverse (http://mathworld.wolfram.com/Pseudoinverse.html) at MathWorld.*
- *Eric W. Weisstein, Moore-Penrose Inverse (http://mathworld.wolfram.com/Moore-PenroseMatrixInverse.html) at MathWorld.*

Retrieved from "http://en.wikipedia.org/wiki/Moore-Penrose_pseudoinverse"
Categories: Matrix theory | Singular value decomposition | Numerical linear algebra

IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 10, NO. 3, SEPTEMBER 1991

Selection of a Convolution Function for Fourier Inversion Using Gridding

John I. Jackson, Craig H. Meyer, Dwight G. Nishimura, *Member, IEEE*, and Albert Macovski, *Fellow, IEEE*

Abstract—**In fields ranging from radio astronomy to magnetic resonance imaging, Fourier inversion of data not falling on a Cartesian grid has been a problem. As a result, multiple algorithms have been created for reconstructing images from nonuniform frequency samples. In the technique known as gridding, the data samples are weighted for sampling density and convolved with a finite kernel, then resampled on a grid preparatory to a fast Fourier transform. This paper compares the utility of several convolution functions, including one that outperforms the "optimal" prolate spheroidal wave function in some situations.**

I. Introduction

IN FIELDS ranging from radio astronomy to computerized tomography and magnetic resonance imaging, spatial frequency data is used to generate images. In order to take advantage of the great computational speed afforded by the fast Fourier transform [1], the data must lie on a Cartesian grid. However, because of hardware constraints or practical considerations, this is not always feasible. Accordingly, many algorithms have been developed for reconstruction from nonuniform samples. Some methods use various interpolation schemes, such as nearest-neighbor, bilinear interpolation, and truncated sinc function finite impulse response interpolators [2]–[4]. Other techniques include gradient descent methods [5], or reconstruction using coordinate transformation [6]. For data on a polar grid, such as may be encountered in computerized tomography or diffraction tomography, filtered back-projection [7] may be used for reconstruction.

In this paper we consider the algorithm known as "gridding." In its earliest form, as first used by radio astronomers, the spatial frequency plane was divided into a grid, and the point in the center of each "cell" was then assigned a value equal to the sum of all of the data points falling within the grid [8]. Later improvements included the use of the average value of the data within the cell [9], or weighting the sampling points based on the distance from array points, such as with a Gaussian function [10]. To account for variations in the spectral sampling density, a further modification of the Gaussian weighting method normalizes the sum of the coefficients to be applied to the sampling points to unity. These cell summing, cell averaging, and Gaussian methods have been compared by Thompson and Bracewell [11].

An overview of the gridding operation is given by O'Sullivan [12], who shows that the optimal gridding method is convolution with a sinc ($\sin \pi x/\pi x$) function of infinite extent, followed by sampling onto a Cartesian grid. Practical considerations require that the infinite sinc function be replaced with a finite convolving function. This paper compares the artifact introduced into the image for various convolving functions of different sizes, including the Kaiser–Bessel window and the zero-order prolate spheroidal wave function (PSWF). We also show a convolving function that improves upon the PSWF in some circumstances.

II. Gridding Algorithm

Consider a two-dimensional function $m(x, y)$ with Fourier transform $M(u, v)$ given by

$$M(u, v) = \int_{-\infty}^{\infty} m(x, y) \exp\left[-2\pi i(ux + vy)\right] dx\, dy, \tag{1}$$

and a sampling function S consisting of P two-dimensional delta functions at positions u_j, v_j,

$$S(u, v) = \sum_{j=1}^{P} {}^2\delta(u - u_j, v - v_j). \tag{2}$$

The sampled Fourier data is given by

$$M_S(u, v) = M(u, v) \cdot S(u, v). \tag{3}$$

In gridding, the sampled data is convolved with a function $C(u, v)$ [such as a Gaussian, a sinc, or a small finite window] and sampled onto a unit spaced grid,

$$\begin{aligned} M_{SCS}(u, v) &= [M_S(u, v) * C(u, v)] \cdot \text{Ш}(u, v) \\ &= \{[M \cdot S] * C\} \cdot \text{Ш} \end{aligned} \tag{4}$$

where * denotes two-dimensional convolution, and the shah or comb function $\text{Ш}(u, v)$ is defined as a sum of equally spaced two-dimensional delta functions:

$$\text{Ш}(u, v) = \sum_i \sum_j {}^2\delta(u - i, v - j). \tag{5}$$

Manuscript received November 16, 1990; revised May 8, 1991. This work was supported by the GE Medical Systems Group; the National Institutes of Health under Grants HL-34962, HL-39478, and HL-39297; and the National Cancer Institute under Grants CA 50948 and CA 48269.

The authors are with the Magnetic Resonance Systems Research Laboratory, Stanford University, Stanford, CA 94305.

IEEE Log Number 9102132.

The corresponding reconstructed image m_{SCS} is given by the inverse Fourier transform of M_{SCS},

$$m_{SCS}(x, y) = \{[m(x, y) * s(x, y)] \cdot c(x, y)\} * \text{III}(x, y). \tag{6}$$

Here we can observe the effect of $S(u, v)$ on the reconstructed image. First, the inverse Fourier transform of the original sampling function $S(u, v)$, which we refer to as $s(x, y)$, affects the aliasing of $m(x, y)$ at a level that cannot be recovered. Thus, as expected, if the function $M(u, v)$ is not sufficiently sampled, the aliasing cannot be corrected via postprocessing. We can, however, make a correction for a nonuniform sampling density in $S(u, v)$ by introducing an *area density function* $\rho(u, v)$. By defining

$$\rho(u, v) = S(u, v) * C(u, v), \tag{7}$$

areas that are oversampled will have a large area density, while areas that are undersampled will have a small area density. As discussed by Bracewell and Thompson [13], the *principal transfer function* (PTF) is given by $S(u, v)/\rho(u, v)$, and the *principal response pattern* by the two-dimensional Fourier transform of the PTF. Introducing the area density function into the reconstruction, we generate the sampled, weighted, convolved, and sampled M,

$$M_{SWCS}(u, v) = \left\{\left[\frac{M_S(u, v)}{\rho(u, v)}\right] * C(u, v)\right\} \cdot \text{III}(u, v),$$

$$= \left(\left\{M \cdot \left[\frac{S}{S * C}\right]\right\} * C\right) \cdot \text{III}. \tag{8}$$

Note that the weighted and convolved M_S can be considered a moving weighted average where the sum of the weighting coefficients as determined by $C(u, v)$ has been normalized to one. The corresponding image is given by

$$m_{SWCS}(x, y) = (\{m(x, y) * [s^{*-1}(s \cdot c)]\} \cdot c) * \text{III} \tag{9}$$

where $*^{-1}$ refers to a deconvolution.

We now consider the effect of $C(u, v)$ on the reconstructed image. As noted by O'Sullivan [12], the optimal convolution function is an infinite sinc, but this function is computationally impractical. A finite convolving function will contribute sidelobes, which will be aliased back into the image by the shah function. Also, any rolloff in the central lobe of $c(x, y)$ will show up as an attenuation towards the sides of the image. This rolloff can be corrected by dividing by $c(x, y)$. This flattens the response across the central lobe, but amplifies the effective amplitude of the aliasing sidelobes, as shown in Fig. 1 for a Hamming window. As is usually done, we will also limit the image to a finite region of interest containing only one "replication" of the object. This is represented mathematically by multiplying by a two-dimensional rect or boxcar function ${}^2\sqcap(x, y)$ where

$${}^2\sqcap(x, y) = \begin{cases} 1 & \text{if } |x| < 0.5 \text{ and } |y| < 0.5 \\ 0 & \text{otherwise} \end{cases}. \tag{10}$$

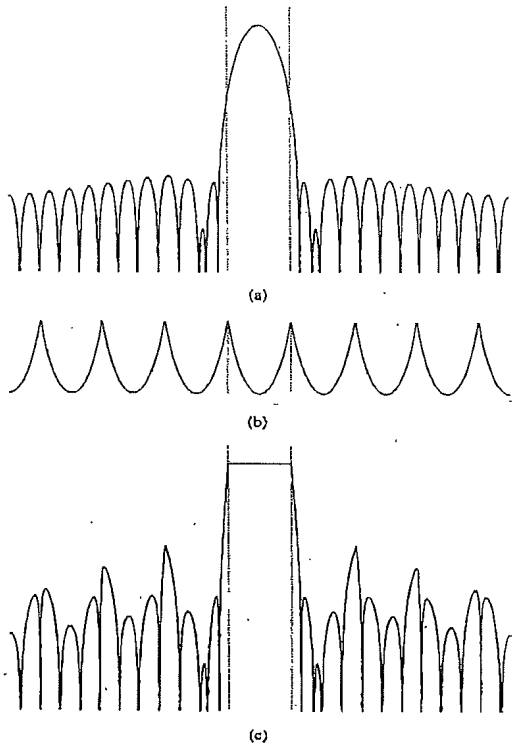

Fig. 1. (a) The log-scale inverse Fourier transform of a Hamming window. The dotted lines represent the edges of the image, with energy in the inverse Fourier transform outside of the dotted lines aliasing back into the image. (b) The convolution rolloff correction for the image. (c) The effective inverse Fourier transform after multiplying by the rolloff correction.

We call the generated image m^*,

$$m^*(x, y) = m_{SWCS} \cdot \frac{{}^2\sqcap(x, y)}{c(x, y)}$$

$$= [(\{m(x, y) * [s^{*-1}(s \cdot c)]\} \cdot c) * \text{III}] \cdot \frac{{}^2\sqcap}{c}, \tag{11}$$

and we refer to the method of generating this image as the *gridding algorithm*. The process is illustrated in Fig. 2.

III. Convolution Function Comparison

The utility of any convolution function is determined by the amplitude and placement of the aliasing sidelobes *after* the image has been corrected for rolloff near the edges. For example, gridding using a $\sqcap(x)$ function will suffer from large aliased sidelobes in the image, but none of the sidelobes will alias into the center of the image. If

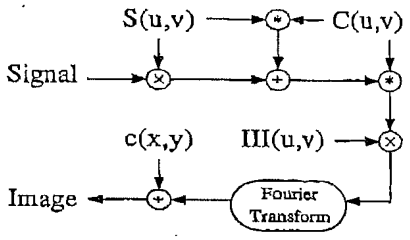

Fig. 2. The gridding algorithm. The signal is sampled by a function $S(u, v)$. The resulting data samples are convolved by the function $C(u, v)$, which is also used to generate the area sampling density weighting. The convolved data are sampled onto a Cartesian grid by multiplying by $\mathrm{III}(u, v)$. Finally, the inverse Fourier transformed data are divided by $c(x, y)$ to compensate for rolloff in the inverse Fourier transform of $C(u, v)$.

the inverse Fourier transform of the convolution function rolls off too much within the bounds of the image, the aliasing into the sides of the image will be greatly amplified when the rolloff is compensated.

We will compare one-dimensional inverse Fourier transforms from separable functions, so $C(u, v) = C(u)\,C(v)$, and $c(x, y) = c(x)\cdot c(y)$. For convenience in working with scale factors, we define the width of the desired image to be 1 unit. The inverse of this, 1 unit of spatial frequency, is the spacing between sampling points in the shah function.

The performance measure that we will use is the relative amount of aliasing energy, including the effect of the rolloff correction. The corresponding functional to be minimized is

$$J = \frac{\displaystyle\int_{|x|>0.5} \left| c(x) \cdot \left[\frac{\sqcap(x)}{c(x)} * \mathrm{III}(x) \right] \right|^2 dx}{\displaystyle\int_{-\infty}^{\infty} \left| c(x) \cdot \left[\frac{\sqcap(x)}{c(x)} * \mathrm{III}(x) \right] \right|^2 dx}, \tag{12}$$

and this is the measure that we will use for comparison.

If all regions within the image are not of equal interest, the performance measure can include a spatially varying weight function, such as Schwab's $w(x, y) = [1 - (2x)^2]^\alpha [1 - (2y)^2]^\alpha$ (where α is a design parameter) [14], which is included in the functional

$$R = \frac{\displaystyle\iint_A |c(x, y)|^2 w(x, y)\, dx\, dy}{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |c(x, y)|^2 w(x, y)\, dx\, dy} \tag{13}$$

where A is the region of interest. Our equation (12) is similar to (13) where all regions within the image are of equal interest (so $w(x, y) = 1$), except that (12) includes the rolloff correction effect on the image, and we are considering separable convolving functions, which simplifies the problem to one dimension.

We will consider the following convolving functions:

1) two-term cosine,

$$\alpha + (1 - \alpha) \cos\left(\frac{2\pi}{W} u\right),$$

2) three-term cosine,

$$\alpha + \beta \cos\left(\frac{2\pi}{W} u\right) + (1 - \alpha - \beta) \cos\left(\frac{4\pi}{W} u\right),$$

3) Gaussian,

$$\exp\left[-\frac{1}{2}\left(\frac{u}{\sigma}\right)^2\right],$$

4) Kaiser–Bessel,

$$\frac{1}{W} I_0[\beta\sqrt{1 - (2u/W)^2}], \text{ and}$$

5) prolate spheroidal wave function, see [15]–[17].

All functions are defined over $|u| \le W/2$, giving each a width W, with α, β, and σ as free design parameters. The well-known Hamming and Hanning windows are examples of the two-term cosine function (with $\alpha = 0.54$ and $\alpha = 0.50$, respectively), and the Blackman window is an example of the three-term cosine function (with $\alpha = 0.42$ and $\beta = 0.50$) [18]. For a given window width and desired bandwidth B the truncated zero-order PSWF of the first kind contains the least amount of energy outside of the desired passband, i.e., it minimizes

$$\frac{\displaystyle\int_{|x|>B} |c(x)|^2\, dx}{\displaystyle\int_{-\infty}^{\infty} |c(x)|^2\, dx}. \tag{14}$$

The PSWF is quite difficult to compute, but the Kaiser-Bessel function [19] is a good approximation, and both it and its inverse transform are easily calculated. The function itself is based on I_0, the zero-order modified Bessel function of the first kind, and the inverse transform is given by

$$c(x) = \frac{\sin\sqrt{\pi^2 W^2 x^2 - \beta^2}}{\sqrt{\pi^2 W^2 x^2 - \beta^2}} \tag{15}$$

For each of the functions, the parameters α, β, and σ were varied to determine the best possible performance, as measured by J in (12), at each function width. The resulting values of J are shown in Fig. 3 for the parameter values given in Table I. The relatively poor results are because of the difficulty in generating a finite function whose inverse Fourier transform goes immediately from a near unity passband to a very low amplitude stopband, without allowing for a transition band. The consequence is relatively large errors from the aliasing sidelobes that appear near the edges of the image where the division by $c(x)$, to make the passband uniform, also amplifies the corresponding portions of the sidelobes.

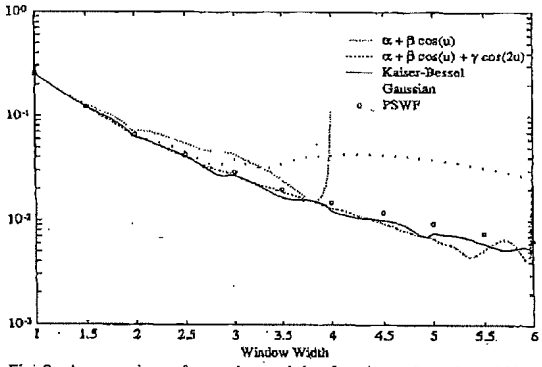

Fig. 3. A comparison of several convolving functions of varying widths, showing the relative amount of unwanted sidelobe energy that will alias into the image if that convolving function is used in the gridding algorithm. Note that when the two-term cosine function reaches a width of four, or when the three-term cosine function reaches a width of six, the first sidelobe of the function's inverse Fourier transform occurs within the image.

TABLE I
THE PARAMETER VALUES FOR EACH FUNCTION TYPE THAT PROVIDE THE LEAST RELATIVE ALIASED ENERGY WHEN GRIDDING ONTO A REGULAR GRID

Window Width	Two-Term cos	Three-Term cos		Gaussian	Kaiser–Bessel
	α	α	β	σ	β
1.5	0.7600	0.8701	0.2311	0.4241	1.9980
2.0	0.7146	0.8099	0.3108	0.4927	2.3934
2.5	0.6185	0.6932	0.4176	0.4839	3.3800
3.0	0.5534	0.5995	0.4675	0.5063	4.2054
3.5	0.5185	0.5383	0.4831	0.5516	4.9107
4.0		0.4998	0.4891	0.5695	5.7567
4.5		0.4653	0.4972	0.5682	6.6291
5.0		0.4463	0.4985	0.5974	7.4302

A simple solution to this problem, alluded to by O'Sullivan [12], is to increase the image field-of-view by sampling onto a smaller grid, and then ignore the outer portion of the image. By doubling the image field-of-view, the central lobe of $c(x)$ can be three times as wide (since it can partially alias around both sides of the image and still not enter the region of interest), which makes possible much smaller amplitude sidelobes [16]. Additionally, $c(x)$ does not taper as much within the region of interest, so less rolloff correction is needed. Conversely, if the former method required Fourier transformation of an $N \times N$ image, subsampling the data in the manner proposed will require transformation of a $2N \times 2N$ image, in addition to computing M_{SWCS} at four times as many points. The computation of M_{SWCS} at each point may be somewhat easier, however, since the subsampling method permits the use of a smaller convolving function for any given error bound.

In addition to the functions previously discussed, we have added a convolving function of our own design. The function was generated in an iterative manner. Starting with any even function, the function was inverse Fourier transformed and those portions of the transformed function that would alias into the region of interest were set to zero. The function was then Fourier transformed back to the original domain and spatially bounded. This process was then repeated many times. This is similar to the method used to generate the PSWF where an even function is repeatedly low-pass filtered and spatially bounded.

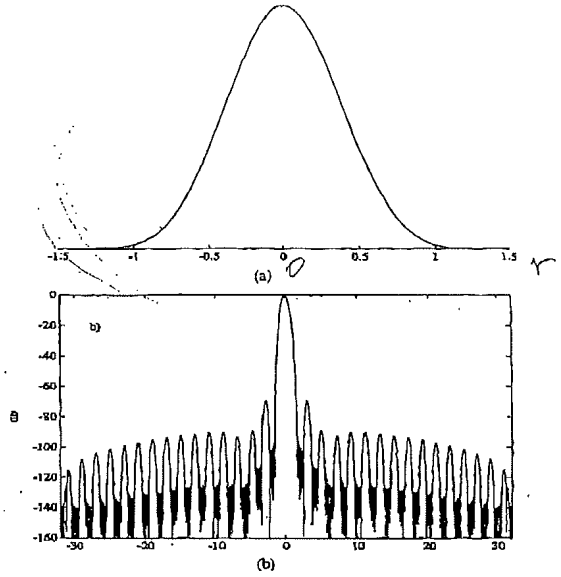

Fig. 4. Authors' width 2.5 convolving function (a) and the corresponding inverse Fourier transform (b). Note the suppression of the portions of the transform that will alias into the center region of the image.

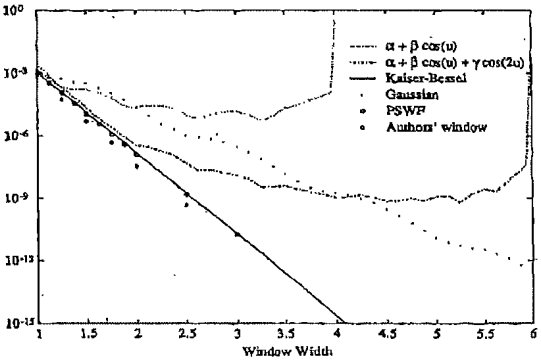

Fig. 5. A comparison of several convolving functions of varying widths, showing the relative amount of unwanted sidelobe energy that will alias into the center region of an image. This center region contains the entire region of interest if the sampled data was gridded onto a 2× subsampled grid.

TABLE II
THE PARAMETER VALUES FOR EACH FUNCTION TYPE THAT PROVIDE THE LEAST RELATIVE ALIASED ENERGY WHEN GRIDDING ONTO A 2× SUBSAMPLED GRID

Window Width	Two-Term cos	Three-Term cos		Gaussian	Kaiser-Bessel
	α	α	β	σ	β
1.5	0.5273	0.4715	0.4917	0.2120	6.6875
2.0	0.5125	0.4149	0.4990	0.2432	9.1375
2.5	0.5076	0.4011	0.4996	0.2691	11.5250
3.0	0.5068	0.3954	0.4997	0.2920	13.9086
3.5	0.5051	0.3897	0.4999	0.3145	16.2734
4.0		0.3850	0.5000	0.3363	18.5547
4.5		0.3833	0.5000	0.3557	
5.0		0.3823	0.5000	0.3737	

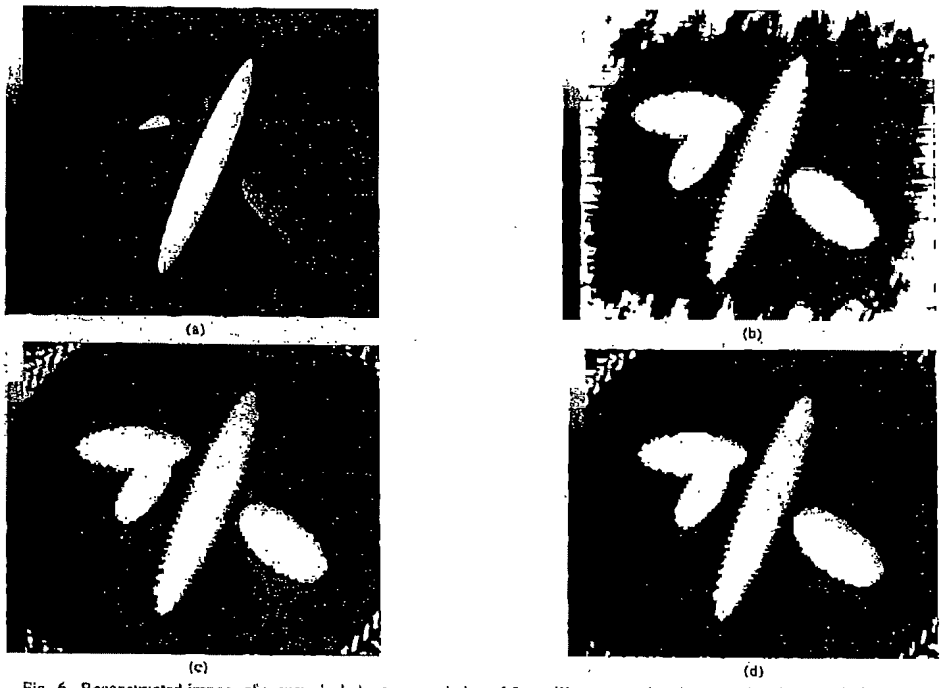

Fig. 6. Reconstructed images of a numerical phantom consisting of four ellipses, showing the effect that the convolution function has on the resulting image. Image (a) was reconstructed using a width three Kaiser-Bessel window (β = 4.2054), without subsampling. Image (b) is the same as image (a), but scaled from 0 to 5% to show the error in the reconstruction. Image (c) was reconstructed using a width three, two-term cosine window [0.5068 + 0.4932 cos ($\pi u/1.5$)] on a subsampled grid, and is scaled the same as image (b). Image (d) was reconstructed using a width three, Kaiser-Bessel window (β = 13.9068) on a subsampled grid, and is also scaled the same as image (b).

The desired PSWF is the eigenfunction of this operation that has the largest eigenvalue, so all undesired portions of the function will gradually decay away relative to the desired function. In Fig. 4 our width 2.5 function is shown. For wider convolving functions, the sidelobes will be increasingly lower, and many more iterations will be required to generate the function. The function shown is the result after 1 000 000 iterations.

As seen in Fig. 5, the reduction in aliased energy is rather dramatic when the gridding is performed on a subsampled grid. The parameters α, β, and σ for each of the function widths are given in Table II.

Fig. 6 shows the effect that the convolving function has on a reconstructed image. Reconstructions were done using a width three Kaiser–Bessel window on a regular grid, a width three two-term cosine window on a subsampled grid, and a width three Kaiser–Bessel window on a subsampled grid. Improvements in the reconstruction error can be seen with each successive technique.

IV. Conclusion

The gridding algorithm can be implemented with virtually no adverse effects from aliasing sidelobes of the convolving function. This requires that the convolved data is sampled finely enough to yield an image field-of-view that is larger than the actual region of interest. We have found that oversampling by a factor of two in each direction is sufficient to yield excellent sidelobe suppression.

The selection of a convolving function requires two primary considerations. First, the performance of the function, and second, the computation time required to generate the function. Although a (discretized) version of a convolving function needs to be computed only once, this can still be a limiting consideration if the generating computation time is extremely long, as is the case with the PSWF and the authors' function. Only slightly poorer results are achieved with the easily computable Kaiser–Bessel window. Use of the Kaiser–Bessel window requires selecting the free parameter, β. The best choice of β, in the sense of minimizing (12), is given in Table II for several window widths.

References

[1] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Mathemat. Comput.*, vol. 19, pp. 297–301, 1965.

[2] S. X. Pan and A. C. Kak, "A computational study of reconstruction algorithms for diffraction tomography: Interpolation versus filtered backpropagation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1262–1275, 1983.

[3] H. Stark, J. W. Woods, I. Paul, and R. Hingorani, "Direct Fourier reconstruction in computer tomography," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 237–245, 1981.

[4] H. Stark, J. W. Woods, I. Paul, and R. Hingorani, "An investigation of computerized tomography by direct Fourier inversion and optimum interpolation," *IEEE Trans. Biomed. Eng.*, vol. BME-28, pp. 496–505, 1981.

[5] W. E. L. Grimson, "A computational theory of visual surface interpolation," *Philosoph. Trans. Roy. Soc. London, Series B*, vol. 298, pp. 395–427, 1982.

[6] J. Clark, M. Palmer, and P. Lawrence, "A transformation method for the reconstruction of functions from nonuniformly spaced samples," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1151–1165, 1985.

[7] R. N. Bracewell and A. C. Riddle, "Inversion of fan-beam scans in radio astronomy," *Astrophys. J.*, vol. 150, no. 2, pp. 427–434, 1967.

[8] N. C. Mathur, "A pseudodynamic programming technique for the design ov correlator supersynthesis arrays," *Radio Sci.*, vol. 4, no. 3, pp.235–244, 1969.

[9] D. E. Hogg, G. H. MacDonald, R. G. Conway, and C. M. Wade, "Synthesis of brightness distribution in radio sources," *Astronom. J.*, vol. 74, no. 10, pp. 1206–1213, 1969.

[10] W. N. Brouw, "Aperture synthesis," in *Methods in Computational Physics, Volume 14*, B. Alder, S. Fernbach, and M. Rotenberg, Eds. New York: Academic, 1975, pp. 131–175.

[11] A. R. Thompson and R. N. Bracewell, "Interpolation and Fourier transformation of fringe visibilities," *Astronom. J.*, vol. 79, no. 1, pp. 11–24, 1974.

[12] J. O'Sullivan, "A fast sinc function gridding algorithm for Fourier inversion in computer tomography," *IEEE Trans. Med. Imaging*, vol. MI-4, pp. 200–207, 1985.

[13] R. N. Bracewell and A. R. Thompson, "The main beam and ringlobes of an east-west rotation-synthesis array," *Astrophys. J.*, vol. 182, no. 1, pp. 77–94, 1973.

[14] F. R. Schwab, "Optimal gridding of visibility data in radio interferometry," in *Indirect Imaging*, J. A. Roberts, Ed. New York: Cambridge University Press, 1983.

[15] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—i," *Bell Syst. Tech. J.*, vol. 40, pp. 43–63, 1961.

[16] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—ii," *Bell Syst. Tech. J.*, vol. 40, pp. 65–84, 1961.

[17] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—iv: Extensions to many dimensions; generalized prolate spheroidal functions," *Bell Syst. Tech. J.*, vol. 43, pp. 3009–3057, 1964.

[18] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, pp. 51–83, 1978.

[19] J. F. Kaiser, *Digital filters*, in *System Analysis by Digital Computer*, F. F. Kuo and J. F. Kaiser, Eds. New York: Wiley, 1966, ch. 7.